Exponent

40276078
Superfund

Technical Memorandum
June 2004

Additional Statistical Analysis of EPA's Soil Recontamination Data

Introduction

This technical memorandum provides additional statistical analysis of the U.S. Environmental Protection Agency's (EPA's) Soil Recontamination Data collected at five properties located within one-quarter mile of the Doe Run smelter in Herculaneum, Missouri.[1] Also presented is a critique of the Tetra Tech EM Inc. (Tetra Tech) report titled, "Statistical Analysis of Lead Samples Collected from Properties Located within a Quarter Mile of the Herculaneum Lead Smelter Boundary" (Tetra Tech 2004). The main objectives of the Tetra Tech report were to 1) evaluate the appropriateness of combining data collected from individual quadrants to estimate the rate of change in lead concentrations over time, and 2) provide estimates of the potential rates of change in lead concentrations, as well as an assessment of uncertainties associated with these estimates. Our analysis focused on these same two objectives.

Summary and Conclusions

At four of the five properties sampled by EPA and located within one-quarter mile of the Herculaneum smelter, soil lead concentrations in at least one quadrant differed significantly from concentrations detected in other quadrants. These differences may be due to localized sources of lead such as streets and gravel driveways, localized disturbances in the soil profile, and/or sampling artifacts introduced by surface heterogeneities and sampling only the top 1/8 to 1/4 inch of soil. These results suggest that it is inappropriate to combine soil lead concentrations from all four quadrants at these properties for the purpose of estimating trends.

In a prior technical memorandum, Exponent reported significant increasing trends in soil lead concentrations in three of four quadrants at Properties 5, 20, and 22, and in quadrant 2 at Property 6. The magnitude of these trends was determined using Sen's nonparametric technique for estimating slope. Slope estimates for individual quadrants where a significant increasing trend was detected range from 1.6 to 10.7 mg/kg per month. Because these slope estimates are based on soil samples collected from the top 1/8 to 1/4 inch of soil (Miller 2004, pers. comm.), they should be reduced by about a factor of six to reflect the soil lead concentration that is present in the top 1 inch of soil, which is the soil sampling depth used for risk assessment purposes and for setting residential lead cleanup levels (U.S. EPA 2003). After correcting for

---

[1] In a previous technical memorandum, we characterized these same five properties as being located within one-third mile of the smelter based on distances reported in EPA's Soil Recontamination Sampling spreadsheet (March 2004 update). However, in the most recent version of this spreadsheet, May 2004, the distances reported for these five properties have been reduced to one-quarter mile or less.

sample depth,[2] slope estimates range from 0.26 to 1.8 mg/kg per month. Assuming an initial background concentration of 50 mg/kg, these data suggest that it will take 16 to 112 years (mean value of 44 years) to reach an average concentration of 400 mg/kg lead in the top 1 inch of soil. Because there are considerable uncertainties in these time estimates, they should not be taken as definitive times at which an average value of 400 mg/kg will be reached at any property, but rather as an indication that the time required for this to occur will be considerable (i.e., on the order of 40 years).

Statistical Analyses

In what appears to be an arbitrary grouping of quadrants, Tetra Tech compared average soil lead concentrations in front yards (quadrants 1 and 2) with average concentrations in back yards (quadrants 3 and 4) using a paired-difference approach and found "no consistent spatial pattern among properties." Average front-yard concentrations were greater than average back-yard concentrations at Properties 22 and 24; back-yard concentrations exceeded front-yard concentrations at Property 5; and no difference was detected at Properties 6 and 20. Because no specific reason is given for performing this test, it is difficult to interpret these results. However, the data do suggest that it is inappropriate to combine data from individual quadrants at properties 5, 22, and 24. To evaluate whether it would be appropriate to group quadrants at a given property for the purpose of statistical analyses, Exponent used Friedman's test.

Friedman's Test

Friedman's test is a nonparametric technique for comparing multiple related sample populations and is described in Gilbert (1987) and Zar (1984). The technique operates on the relative concentrations of the data, does not require the data to be normally distributed, and can accommodate non-detect values. The technique is used to test the null hypothesis that no one population has larger or smaller values than any of the other populations. In this case, Friedman's test was used to determine whether lead concentrations are similar in each of the four quadrants at a given property.

The test is conducted by ranking the soil lead concentrations in each quadrant during a given sampling event from lowest to highest and then summing the ranks in each quadrant over all sampling events. An example is presented below for Property 20. Soil lead concentrations are presented in Table 1, and relative ranks are presented in Table 2.

[2] As a first approximation, slope estimates were corrected for sample depth by dividing by six.

Table 1. Soil lead concentrations detected at Property 20

Sampling Event	Soil Lead Concentration (mg/kg)			
	Quadrant 1	Quadrant 2	Quadrant 3	Quadrant 4
7	60	57	56	121
8	67	71	69	138
9	86	90	126	61
10	74	84	125	113
12	121	164	152	140
13	150	88	179	137
14	245	210	212	132
15	224	128.5	181.5	295

Table 2. Relative ranks of soil lead concentrations detected for each sampling event at Property 20

Sampling Event	Soil Lead Concentration			
	Quadrant 1	Quadrant 2	Quadrant 3	Quadrant 4
7	3	2	1	4
8	1	3	2	4
9	2	3	4	1
10	1	2	4	3
12	1	4	3	2
13	3	1	4	2
14	4	2	3	1
15	3	1	2	4
Sum of Ranks	18	18	23	21

The Friedman test statistic (Fr) is calculated according to equation (1).

$$F_r = \frac{12}{nk(k+1)} \sum_{j=1}^{k} R_j^2 - 3n(k+1) \qquad (1)$$

where:

n	=	number of sampling events
k	=	number of sample populations
R_j	=	sum of the ranks for population j.

At Property 20, the Friedman test statistic is calculated to be 1.35 using equation (1). This value is compared to the chi-square statistic at the 95% probability level and k-1 degrees of freedom ($\chi^2_{0.95,3} = 7.81$) to determine whether we can reject our null hypothesis of no difference between soil lead concentrations in each quadrant. Because the Friedman's test statistic of 1.35 is less than the chi-square value of 7.81, we cannot reject our null hypothesis and must conclude that there is no difference between quadrants. Friedman's test results for other properties located within one-quarter mile of the smelter are reported in Table 3.

Table 3. Friedman's test results for properties located within one-quarter mile of the smelter

Property ID	Sum of Ranks				Friedman's Test Statistic[a]
	Quadrant 1	Quadrant 2	Quadrant 3	Quadrant 4	
5	14.5	13.5	24.5	27.5	**11.46**[b]
6	21	20	27	12	**8.55**
20	18	18	23	21	1.35
22	19	28	11	12	**15.86**
24	24	12	11	13	**11.00**

[a] Friedman's test statistic was calculated using a modified version of equation (1) that involves a correction for ties (Gilbert 1987).

[b] Bold typeface indicates a significant difference among quadrants.

At Properties 5, 6, 22, and 24, the null hypothesis of no difference between soil lead concentrations in individual quadrants is rejected using Friedman's test. Table 3 includes the sum of ranks for each quadrant to give an indication of which quadrants are different. Zar (1984) provides a multiple comparison technique for identifying differences between populations if the null hypothesis is rejected. The difference between the sum of ranks for any two populations is divided by the standard error [given by equation (2)] and compared to a table of critical values.

$$SE = \sqrt{\frac{nk(k+1)}{12}} \tag{2}$$

For example, at Property 22, quadrant 2 appears to be different from quadrants 3 and 4, and possibly quadrant 1. Comparing quadrants 2 and 4, the difference between the sum of ranks is 16; the standard error for $n = 7$ and $k = 4$ is 3.42; and the test statistic is $16 / 3.42 = 4.68$. The critical value for comparison is 3.63 (obtained from a lookup table in Zar). Thus, we can conclude that soil lead concentrations in quadrant 2 are significantly different from (higher than) soil lead concentrations in quadrants 3 and 4. For quadrants 1 and 2, the test statistic is $9 / 3.42 = 2.63$, and we cannot conclude that soil lead concentrations are significantly different.

At Property 5, soil lead concentrations in quadrant 4 are greater than concentrations in quadrant 2 but not quadrants 1 and 3; at Property 6, soil lead concentrations in quadrant 3 are greater than concentrations in quadrant 4 but not quadrants 1 and 2; and at Property 24, soil lead concentrations in quadrant 1 are greater than concentrations in quadrants 2 and 3 but not 4.[3] These results suggest that a localized source (such as the redistribution of lead from a nearby street or gravel driveway) or a localized disturbance of soil may be preferentially affecting lead concentrations in one of the quadrants. The results also suggest that it is inappropriate to combine soil lead concentrations from all four quadrants at these properties for the purpose of estimating trends.

Tetra Tech compared average soil lead concentrations detected in the front yard (quadrants 1 and 2) with average soil lead concentrations detected in the back yard (quadrants 3 and 4) and found "no consistent spatial pattern among properties." No reason is given for grouping the data in this fashion, and the authors admit that their results are "not surprising, as these groupings are not based on absolute orientations of front or backyards (or in effect, the relative degree of exposure of each area to air emissions of lead)."

It is unclear how this test, whatever the outcome, provides information to make a determination about "the most appropriate way to use data collected from individual quadrants in any statistical analysis of temporal trends," which is one of the stated objectives in the report. The authors suggest that grouping quadrants as "exposed" or "relatively unexposed" is a "more technically defensible" approach, but it is unclear how such designations would be made, and what the results would mean.

Slope Estimates

Tetra Tech used linear regression analysis to obtain slope estimates at Property 20 and for all properties combined. However, linear regression is not the best statistical tool for this use. Linear regression analysis is sensitive to extreme values and does not easily handle censored data. It also requires assumptions that the data are normally distributed and have constant variance, which may not be true for the soil recontamination data. Because of these limitations, EPA does not recommend using linear regression as a general tool for estimating and detecting trends (U.S. EPA 2000). This is the same conclusion drawn by EPA's statistician after

[3] *The test statistic for comparing quadrants 1 and 4 is 3.48, which is slightly less than the critical value of 3.63.*

reviewing the soil recontamination data at this site (Wolf undated; Wolf 2004).[4] For analysis of the soil recontamination data, Sen's method for estimating slopes is a more robust tool. This method was used for estimating slopes at properties with significant trends in one or more quadrants.

Sen's Estimate of Slope

Sen's method is a nonparametric technique for estimating a linear slope for a time series of data that exhibit a trend (Gilbert 1987; U.S. EPA 2000). For a time series of n data points, slope estimates are calculated for all possible n (n-1) / 2 data pairs, and the median slope estimate is taken as the overall slope. The first step is to convert each sampling event to an elapsed time, in days, with day zero beginning on 8/26/2002 (sampling event 7) as shown in Table 4, using soil lead concentrations detected in quadrant 1 of Property 20 as an example.

Table 4. Soil lead concentrations detected at Property 20, quadrant 1

Sampling Event	Date	Elapsed Time (days)	Soil Lead Concentration (mg/kg)
7	8/26/2002	0	60
8	9/24/2002	29	67
9	11/7/2002	73	86
10	12/10/2002	106	74
12	3/17/2003	203	121
13	6/23/2003	301	150
14	9/23/2003	393	245
15	12/22/2003	483	224

With eight sampling events, there are 28 possible slope estimates. For example, using sampling events 7 and 8, the estimated slope is (67 – 60)/(29 – 0) = 0.24 mg/kg-day; using events 7 and 9, the estimated slope is (86 – 60)/(73 – 0) = 0.36 mg/kg-day; using events 7 and 10, the slope is (74 – 60)/(106 – 0) = 0.13 mg/kg-day, etc. After all 28 possible slope estimates are calculated, the median value is selected as the estimate of the true slope. For Property 20, quadrant 1, the median slope is 0.34 mg/kg-day or 10.3 mg/kg per month.

Friedman's test indicates that the four quadrants at Property 20 are not significantly different from each other.[5] Thus, it makes sense to combine quadrants and calculate an overall slope for

[4] Memos are attached to this report.

[5] Note that significant increasing trends were detected using the Mann-Kendall test in quadrants 1, 2, and 3 at Property 20 but not in quadrant 4.

the property. With eight samples in each of four quadrants, there are 32(31)/2 = 496 data pairs; however, samples collected in the same sampling event cannot be compared. This leaves 448 possible data pairs. The overall median slope of the 448 data pairs at Property 20 is 0.30 mg/kg-day, with 95% confidence limits of 0.294–0.312 mg/kg-day.[6] This translates to 9.1 mg/kg per month (95% confidence limits of 8.9 to 9.5 mg/kg per month). Slope estimates for individual quadrants showing a significant increasing trend based on the Mann-Kendall tests are reported in Table 5.

Table 5. Sen's slope estimates for quadrants that show a significant increasing trend using the Mann-Kendall test

Property ID	Sen's Estimate of Slope (mg/kg per month)			
	Quadrant 1	Quadrant 2	Quadrant 3	Quadrant 4
5	2.3	1.6	3.5	--
6	--	4.6	--	--
20	10.3	7.6	9.4	--
22	--	10.7	4.8	3.1
24	--	--	--	--

-- No significant trend detected using the Mann-Kendall test.

Tetra Tech reported an uncorrected slope of 22 mg/kg per month for Property 20 using Sen's technique. This estimate appears to be in error; it differs significantly from our estimate (calculated by the same method), and is almost three times Tetra Tech's estimate obtained using linear regression analysis. Tetra Tech also reported a slope of 8 mg/kg per month for all properties combined using Sen's technique. This estimate is also likely to be in error. In addition, there is no statistical basis for estimating a slope at properties where no significant trend in soil lead concentrations has been detected.

Interpreting Slope Estimates from EPA Soil Recontamination Data

At a March 16, 2004 meeting called by the City of Herculaneum, Bruce Morrison of EPA described the soil sampling method used to collect EPA's Soil Recontamination data as "scraping the top 1/8 to 1/4 inch of the soil" (Miller 2004, pers. comm.). This is in contrast to the Quality Assurance Project Plan for the Site, which specifies collecting composite samples from the upper 1 inch of soil (U.S. EPA 2001). Because the proposed source of soil recontamination is smelter fallout onto the soil surface, soil lead concentrations measured in the

[6] Confidence limits were calculated using the procedure described on page 219 in Gilbert (1987).

top 1/8 to 1/4 inch of soil could greatly overestimate average concentrations in the top 1 inch of soil, particularly at levels significantly above background. The Superfund Lead-Contaminated Residential Sites Handbook (U.S. EPA 2003) states on page 26: "With respect to risk assessment, the top inch of soil best represents current exposure to contaminants and is the source of data used in the IEUBK model to represent exposure from soil." Thus, the top 1 inch of soil is used for establishing cleanup levels for lead in residential soils, and should be used for evaluating trends if the magnitude of the trends will be used to estimate the time to reach cleanup levels.

Corrected slope estimates were calculated by dividing the values presented in Table 5 by a factor of six—in effect, assuming an average sample depth of 1/6 of an inch. These corrected slope estimates are presented in Table 6.

Table 6. Sen's slope estimates after correcting for sample depth for quadrants that show a significant increasing trend using the Mann-Kendall test

Property ID	Sen's Estimate of Slope (mg/kg per month)			
	Quadrant 1	Quadrant 2	Quadrant 3	Quadrant 4
5	0.39	0.26	0.58	--
6	--	0.77	--	--
20	1.7	1.3	1.6	--
22	--	1.8	0.81	0.52
24	--	--	--	--

-- No significant trend detected using the Mann-Kendall test.

Time estimates vary from 16 years for soil lead concentrations in quadrant 2 at Property 22 to 112 years for soil lead concentrations in quadrant 2 at property 5, with a mean value of 44 years for all quadrants with increasing trends. Friedman's test shows that quadrant 2 at Property 22, where the greatest slope is observed, is significantly higher than other quadrants on the property. This suggests that a local source of lead, such as redistribution from a nearby street, driveway, or alleyway, may be affecting this quadrant. This is consistent with the observation that none of the properties show consistent trends across all quadrants, just as adjacent properties show dissimilar soil lead concentration trends. Overall, these results suggest that sources other than smelter emissions are contributing to the observed increases in soil lead concentrations.

References

Gilbert, R.O. 1987. Statistical methods for environmental pollution monitoring. Van Nostrand Reinhold, New York.

Miller, A. 2004. Personal communication (telephone conversation with Aaron Miller on May 3, 2004 regarding the public meeting held March 16, 2004 in Herculaneum Missouri). Doe Run Company.

Tetra Tech. 2004. Statistical analysis of lead samples collected from properties located within a quarter mile of the smelter boundary Herculaneum lead smelter. Prepared for U.S. Environmental Protection Agency, Region 7, Kansas City, KS. Tetra Tech EM Inc., Lenexa, KS.

U.S. EPA. 2000. Guidance for data quality assessment – Practical methods for data analysis. EPA QA/G-9, QA00 Update. EPA/600/R-96/084. U.S. Environmental Protection Agency, Office of Environmental Information. July.

U.S. EPA. 2001. Quality assurance project plan for a site characterization at the Herculaneum lead smelter – Herculaneum, Missouri. CERCLIS ID No. MOD006266373. U.S. Environmental Protection Agency. September.

U.S. EPA. 2003. Superfund lead-contaminated residential sites handbook. OSWER 9285.7-50. U.S. Environmental Protection Agency, Office of Emergency and Remedial Response. August.

Wolf, M. Undated. Memo to Bruce Morrison regarding analyses of lead smelter data.

Wolf, M. 2004. Follow-up e-mail to Bruce Morrison regarding the use of regression analysis and the Mann Kendall test to analyze the lead smelter data. April 1.

Zar, J.H. 1984. Biostatistical analysis. Prentice-Hall, Inc., Englewood Cliffs, NJ.

Michele Wolf
<mwolf@neptuneinc.or g>

04/01/2004 01:54 PM

To: Bruce Morrison/SUPR/R7/USEPA/US@EPA
cc: Kelly Black <kblack@neptuneinc.org>
Subject: Follow-up Discussion

Bruce,

I wanted to follow up with a short email elaborating a bit on the two statistical tests, Regression Analysis and Mann-Kendall, used to analyze the lead smelter data. There are several assumptions that need to be satisfied before using Regression analysis and making inferences from that analysis. Those assumptions are that the data are independent and randomly collected, that there is a linear association between the x and y variables, that the y values are normally distributed for each x value and that the variance for the y values is constant. If one or more of these assumptions is violated, the results of the analysis may be incorrect or misleading. Mann-Kendall does not make any of these assumptions. Furthermore, Mann-Kendall is not greatly affected by outliers, missing data or censored data. For the lead smelter data, I did not test all of the assumptions and do not know if the statisticians who wrote "Herc Analysis" tested these assumptions, however it is clear that the data are not independent of one another. The other assumptions may be violated as well, however statistical tests would need to be performed to evaluate those assumptions. Therefore, Mann-Kendall is the more appropriate test for trend for this data.

You mentioned you would like a small discussion about House 8. From the "Lead Soil Trend Analysis," Mann-Kendall showed a statistically significant positive trend. Although the graph of the data does not appear to show this trend, a change in the x and/or y axis may show this trend more clearly. While the statistical test indicates the trend is significant, there are two reasons it may not be of practical significance. First, many of the data points are non-detects and second, all of the detected concentrations are still very low as compared to the PRG of 400 mg/kg. I would continue to monitor this house closely and re-evaluate its status after the next round of data collection.

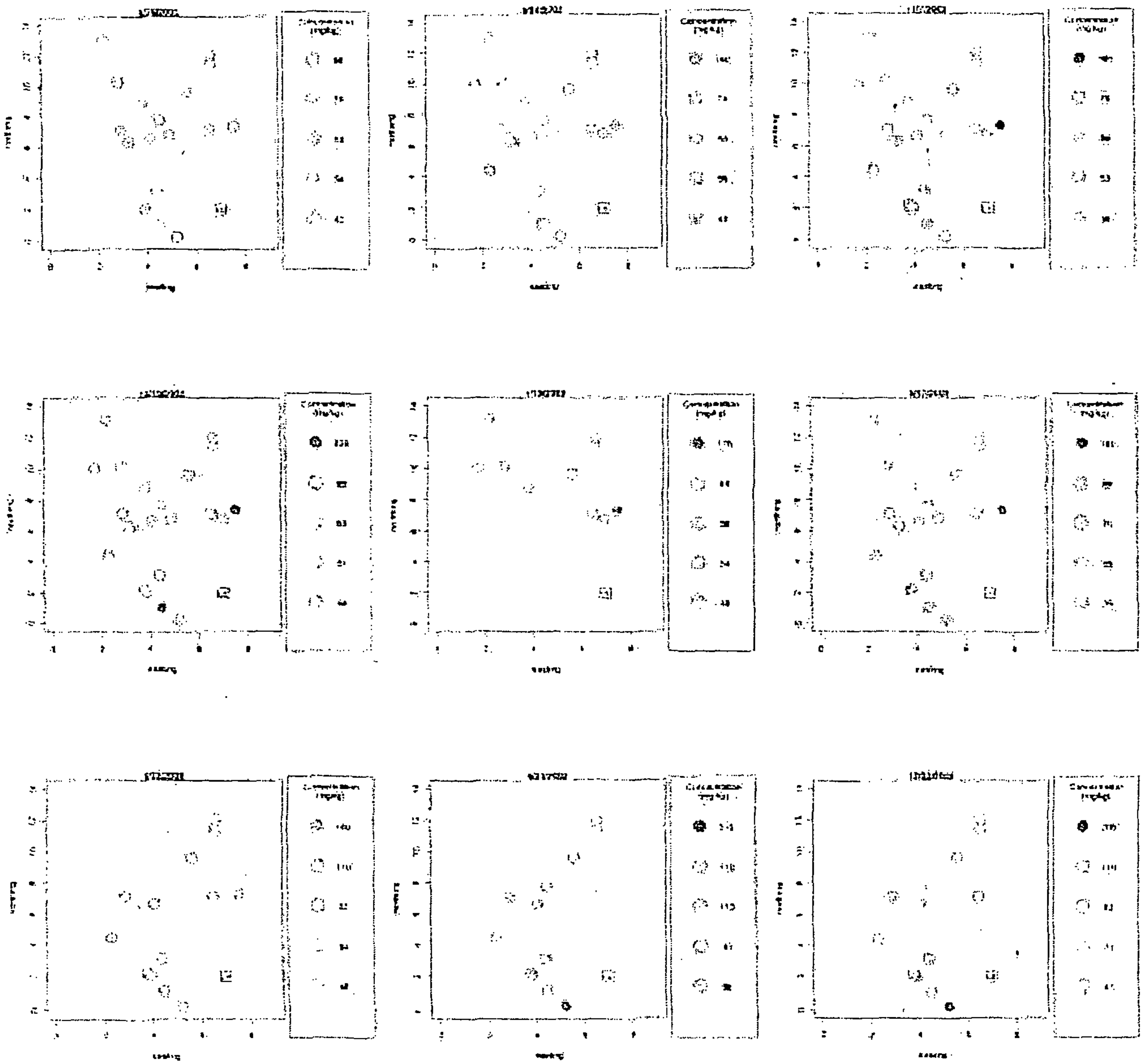
I hope this helps. Please let me know if you have any other questions.

Hope the fish were biting.

Michele

Michele Wolf
Statistician
Neptune and Company
(720)- 746 -1803 Ext. 3
mwolf@neptuneinc.org

Check our website at www.neptuneandco.com



Bruce,

I have had some time to look over both of the documents you sent and perform some of my own exploratory data analyses. The two reports present slightly different statistical approaches to analyzing the data, however, the results are consistent between the two reports. The statisticians who wrote these reports have conducted the same analyses I would have for these data. The Mann-Kendall test is the most appropriate statistical test for these data to evaluate trend. Their graphical output is also the kind I would have presented to visually show the concentrations for each house through time. Summary statistics are usually also informative, however in this case the summary statistics may be misleading since not every house is included in every round.

While both papers present the data by round, I believe it would be better to present the data by date. The time between rounds is not consistent, so visually the trend appears to be increasing sharply in some houses where it may not be so severe if the data were plotted by date rather than by round. The authors of "Herc Analysis" converted the earliest sampling round (Round 7) to zero and all subsequent sampling dates were converted to number of days from this initial sampling date. This method works well to better assess visual trends through time, however they only applied this method to the regression when it should have been applied to both the box plots and the regressions. Finally, some spatial plots showing lead concentrations for each house for each time period might be beneficial for understanding these data. These can be very helpful in determining patterns in space through time.

Since I felt spatial plots were important, I constructed nine intensity plots, one for each time period (each Round). Please find these plots in the word document attached to this email. The concentration for each house is plotted as a circle and filled with a color corresponding to the relative concentration for that house. The colors range from lime green (low concentrations) to red (high concentrations), spanning the continuum from yellow to orange in between. Visually it is clear that, in general, the concentrations increase through time. Furthermore, the houses close to the smelter (denoted as LS) and those downwind are more affected by the smelter than other houses. This can be seen by the changes from lime green to orange/red over time.

Based on the Mann-Kendall analyses, box plots and spatial plots, it is clear that lead concentrations in general are increasing through time. The houses close to the smelter and those downwind appear to be affected more than those further away. It is also clear that there are more statistically significant increasing trends than would be expected due to chance alone. It would be nice to be able to accurately predict future lead concentrations, however, regression analyses for these data are not appropriate for a variety of reasons (EPA QA/G-9S, EPA 2004). Furthermore, extrapolation beyond the range of this data set is not advised. One way to obtain an estimate of the slope to evaluate future lead concentrations would be to use a nonparametric slope estimate. The authors of "Herc Analysis" calculated such an estimate, Sen's slope estimate, and presented a value of 22 (indicating lead concentrations increase 22 mg/kg every month) for the worst-case scenario (House 20). This estimate not only could be used as a conservative estimate of the increasing tendency, but also could be used to predict when lead concentrations will be above the PRG of 400ppm. For example, the median lead concentration for House 20 during Round 15 was 202.75 mg/kg. Using Sen's slope estimate, we would predict that House 20 could be above the PRG within the next nine months. I would reiterate, it *could* be above the PRG within nine months. By using the house with the strongest increasing trend, we get an extremely conservative estimate of trend. Therefore, although the nine month estimate is appropriate for this one house, it is certainly a worst-case guess for any of the other houses based on these data.

It is impossible to predict what will happen in the future, but from these data it is clear that lead concentrations are increasing over time and that some houses will exceed the PRG sometime in the future if the trend continues linearly as is. If operations at the smelter change (either by installing additional controls or by increasing lead emissions), then a change in the trend would be expected. If all conditions remain constant, then the existing trends are the best estimates available for future site conditions. At this point, all we can say is that the number of homes with statistically significantly increasing lead concentrations are beyond the number that would occur by chance alone, and that the rate of increase at the most impacted houses could lead to an exceedance of the lead threshold in the not too distant future.

I hope this analysis has been helpful, please let us know if you have any additional questions or concerns.

Michele Wolf